The rotor fault prediction based on support vector regression and phase space reconstruction

Xiao Han
Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China

ABSTRACT: Support vector regression (SVR) is a popular machine learning method that develops these years and has been widely used in the prediction field. But the input feature vectors largely affect the accuracy of the forecast error, so the feature vector choice has been the hot issues of attention and research scholars. For these problems, some scholars have proposed a characteristics selection method of support vector regression machine based on the phase space reconstruction, but the value of the time delay and embedding dimension became discussion hotspot. For these problems, some scholars have proposed characteristics selection method of the support vector regression machine based on phase space reconstruction, but the value of the time delay and embedding dimension became discussion hotspot. So the optimization method of particle swarm optimization (PSO) is proposed. This method is able to quickly identify the best combination of parameters (τ, m, C, σ) and improve forecast accuracy. This method is applied to the prediction of the rotor misalignment of rotating machinery fault data. The experiment proved that the method is feasible.
KEYWORD: support vector regression; phase-space reconstruction; feature selection; particle swarm optimization; state forecast

1 INTRODUCTION

Today, large rotating machinery such as steam turbine, wind turbine and compressor play a very important role in chemical industry, metallurgy, and power plant. However, as the rotating machinery are constantly moving towards large-scale, complex, high-speed, continuous and automated direction, the demand of the condition monitoring is getting strict. Rotor misalignment is one of the common rotating machinery faults. This fault will cause many damages such as strong vibration of the unit, oil film instability, bearing damage and the deformation of the shaft. So early discovery and early treatment of the rotor fault is one of the effective methods to reduce failure. At present, the most common prediction methods are stochastic time series [1], artificial neural network [1], and neural network [2-3], wavelet neural network [4] chaos and so on. Compared with the above method, the SVR [5-6] is a kind of better fitting and prediction methods.

Now using SVR method to deal with time series is a hot research issue. While the support vector regression machine is using in actual forecast, there are some problems such as feature selection and extraction. General common method is simply characterized by time series for training, but it contains less information, which will make the prediction precision get low.

To solve these problems, some scholars proposed feature extraction method based on phase space reconstruction. The method is put to improve the precision of prediction to a great extent. But in the process of phase space reconstruction, time delay τ and embedding dimension m have great influence on the performance of the phase space reconstruction. According to the above problems, particle swarm algorithm method was proposed to optimize the time delay and embedding dimension. At the same time using the particle swarm algorithm to find the parameters combination (C, σ) of SVR. Using Particle Swarm Optimization (PSO) algorithm to find the optimal combination of the four parameters (τ, m, C, σ) at the same time.

2 SUPPORT VECTOR REGRESSION MACHINE BASED ON PHASE SPACE RECONSTRUCTION

2.1 *Support vector regression machine*

In the 1990s, Vapnik firstly put forward the SVM theory [7]. Different from the neural network, Support vector machine is a kind of machine learning

technology based on structure risk minimization principle, which has the very good generalization performance [8]. The SVM was originally used for classification problems, and then was successful in the regression problem. When the SVM is often named as support vector regression machine (SVR) when it comes to regression and prediction problem Regression problems are similar to classification problems. Give the training set: $T=\{(x_1, y_1), \ldots, (x_n, y_n)\}$.Assumes that the regression function is a linear function of the problem:

$$y = f(x) = wx + b \tag{1}$$

Construct functions for solving optimization problems

$$\min_{\alpha^{(*)}\in \mathbf{R}^{2n}} \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i x_j) + \varepsilon\sum_{i=1}^{n}(\alpha_i^* + \alpha_i) - \sum_{i=1}^{n} y_i(\alpha_i^* - \alpha_i)$$

$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) = 0 \\ 0 \le \alpha_i^*, \alpha_i \le C, \quad i = 1,2,\ldots,n \end{cases} \tag{2}$$

Get the optimal $\boldsymbol{\alpha}^{(*)} = (\alpha_1, \alpha_1^*, \ldots, \alpha_n, \alpha_n^*)^{\mathrm{T}}$ and calculate the ω and b for the regression function:

$$\boldsymbol{\omega} = \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)x_i \tag{3}$$

$$b = y_i - \sum_{j=1}^{n}(\alpha_j^* - \alpha_j)(x_j x_i) + \varepsilon \tag{4}$$

However, many actual problems are nonlinear problems, so we need expand the situation of linear. The kernel function $K(x_i, x_j)$ mapped the data into high-dimensional space, so the optimization problem is transformed into:

$$\min_{\alpha^{(*)}\in R^{2n}} \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i x_j) + \varepsilon\sum_{i=1}^{n}(\alpha_i^* + \alpha_i) - \sum_{i=1}^{n} y_i(\alpha_i^* - \alpha_i) \tag{5}$$

The introduction of kernel function effectively avoid the "dimension disaster". There are many kinds of kernel functions, the use of kernel function will also be different when dealing with various problems. Now there is not a effective method for the choice of kernel function. Radial basis kernel function is generally used kernel function, it is corresponding to infinite dimension space, in which the limited sample data is linear separable. It also have the advantage of little parameters. So, this paper chose the radial basis kernel function:

$$K(x, y) = \exp\left(-\|x - y\|^2 / \sigma^2\right) \tag{6}$$

2.2 *Phase space reconstruction*

Any component evolution of the system is determined by the interaction with other components [9-10]. All these related components information are shielded in the development process of any component. That means it is available to extract and recover the original system rule through some time series data of any a component. For a certain time series: $X=\{x(1), x(2), \ldots, x(n)\}$

The reconstructed phase space can be expressed as:

$$PX = \begin{bmatrix} x(1) & x(2) & & x(i) \\ x(1+\tau) & x(2+\tau) & \cdots & x(2+\tau) \\ \vdots & \vdots & & \\ x(1+(m-1)\tau) & x(2+(m-1)\tau) & & x(i+(m-1)\tau) \end{bmatrix}$$

τ is time delay, m is embedding dimension, and any column $PX(i)$ of the matrix is a phase point.

3 SVR BASED ON PHASE-SPACE

For solving the existing problem of training samples feature selection, this paper puts forward the support vector regression machine based on phase space reconstruction. The method uses the reconstructed phase space to realize character the adaptive choice. Specific model is as follows:

Firstly, for the given series of time data, $X=\{x(1),x(2),\ldots x(n)\}$, use cross-correlation method and Cao method to compute the time delay and embedding dimension, respectively. Then, the time delay τ and embedding dimension m can reconstruct the phase space.

$PX(i) = (x(i), x(i+\tau), \ldots, x(i+(m-1)*\tau))^T$ The m dimension of the column vector $PX(i)$ will be treated as the feature to establish optimization problem,

$$\min_{\alpha^{(*)}\in R^{2n}} \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(PX_i \cdot PX_j) + \varepsilon\sum_{i=1}^{n}(\alpha_i^* + \alpha_i) - \sum_{i=1}^{n} y_i(\alpha_i^* - \alpha_i). \tag{7}$$

Get the optimal solution through Eq.7, and the regression function is

$$f(PX) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)K(PX_i, PX) + b \tag{8}$$

4 PSO THEORY

PSO (Particle Swarm Optimization) algorithm is a kind of evolutionary computation technology, developed by Dr Eberhart and Dr Kennedy [11-12]. PSO algorithm results from the research on the behavior of birds feed, that the most simple and effective way to find food for each bird is to search for the surrounding area of the nearest itself.

PSO algorithm first initializes a group of random particles (random solution), and then the particles are to follow the current optimal particle search in the solution space, through iteration finding the optimal solution. It assumed the position and speed of the d-dimensional search space of particle i are X^i and V^i, and the particles by tracking the optimal solution to update their own two in each iteration. The first optimal solution finding by particles themselves is the individual extreme; the other one is the optimal solution for the entire population at present, namely the global optimal solution gbest. After finding the optimal solution, the particle updates its velocity and location according to the following formula.

$$v_{i,j}(t+1) = w v_{i,j}(t) + c_1 r_1 [p_{i,j} - x_{i,j}(t)] + c_2 r_2 [p_{g,j} - x_{i,j}(t)] \quad (9)$$

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1), j = 1,2....d \quad (10)$$

Where the w is inertia weight factor, c_1 and c_2 are positive learning factors which are random numbers between 0 and 1 uniformly distributed.

5 PARTICLE SWARM OPTIMIZATION ALGORITHM FOR THE OPTIMAL COMBINATION OF PARAMETERS

The calculation methods of the time delay are C-C algorithm, the cross-correlation algorithm; while the calculation methods of embedding dimension includes G-P algorithm, the pseudo adjacent points, the correlation integral algorithm and Cao algorithm. In this paper, PSO algorithm was applied to calculated the optimal combination of the τ and m. Rounding command (round (A)) was applied to each parameter value to avoiding the decimals in the PSO algorithm calculation progress. The main principle is to find the smallest error value of the prediction accuracy corresponding the parameters combination. Average relative error is used as fitness value:

$$y = sum\left(abs(PL\_rest - srest) / srest\right) / L .$$

Where PL_rest and srest are respectively prediction values and actual values. L is the number of the prediction points.

6 EXPERIMENT STUDY

This test was conducted on the Bentley bench to simulate the rotor misalignment fault. The rotor was set to 1200rpm and the sampling frequency was set to 1000HZ in the test progress. The first 200 data points of the signal were taken for the training points, while the rest 50 points of the signal were chosen as the prediction points. Fig. 1 shows the original image of the vibration signal.

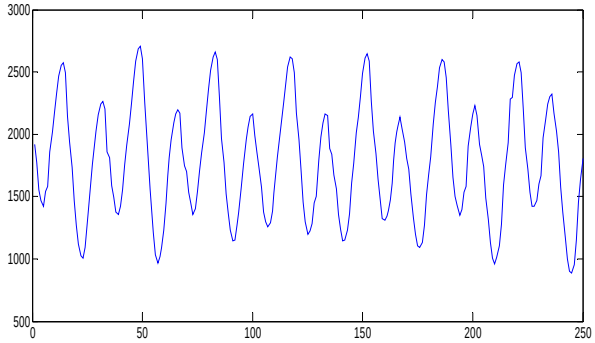

Fig. 1 Experimental vibration signal

The phase space reconstruction of the signal is carried out for the prediction based on the each phase point as the eigenvalues input in the space. RBF kernel function was selected in the SVR. The maximum number of iterations was set 100 and the population size was set 20 respectively in the PSO algorithm. The final results shows that and m are 11 and 4 respectively, and the parameters C=4000, σ=100.

In order to compare the superiority, the cross-correlation algorithm and Cao algorithm were adopted to calculate and m, the value of which are 2 and 8 respectively. The prediction results of the two methods are shown in the Fig. 2 and Fig. 3. The fitting results of Fig. 2 and Fig. 3 show that the method proposed in this paper can attain much better results than the cross-correlation algorithm and Cao algorithm when applied in prediction. Where "+" represent real value and " o "represent prediction value.

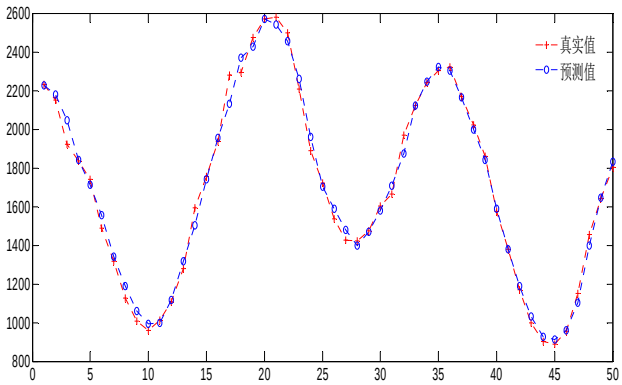

Fig. 2 Prediction results of the proposed

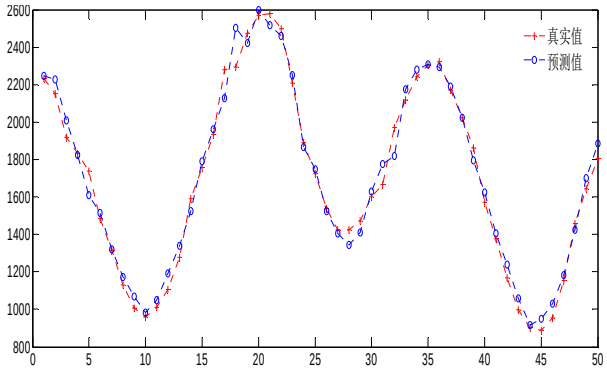

Fig.3 Prediction results of the comparison methods

Table 1 is a comparison of the predicted value between the two methods. Prediction results are evaluated by the mean error (ME) and mean relative error (MRE) shown in the table 1.

Tab. 1 Comparison of predicted results

	Proposed method	Compared method
ME	34.9625	52.9431
MRE (%)	2.19	3.33

As we can be seen from Tab.1, PSO algorithm that used to find the combination τ and m can attain more accurate prediction results compared to the common cross-correlation algorithm and Cao algorithm. Meanwhile, we also get reasonable values of C and σ.

7 CONCLUSION

To solve the selection problem based on the phase space reconstruction of SVR in time delay τ, embedding dimension m and SVR parameters, this paper proposes an optimization method based on PSO. Through with the common method, the results show the superiority of paper method, which improves the prediction accuracy.

REFERENCES

[1] YANG Xiuyuan, XIAO Yang, CHEN Shuyong. Wind speeed and generated power forcasting in wind farm [J]. Proceedings of the CSEE, 2005,25(11):125

[2] Hippert H S, Pedreira C E, Souza R C. Neural networks for short-term load forecasting: A review and evaluation [J]. IEEE Trans on Power System, 2001, 16(1): 44—55

[3] LIU Tao, PENG Shirui. Application of improved BP neural Network to prediction of aviation ammunition consumption [J]. Journal of Detection &Control ,2009, 31(5): 52-55

[4] SHI Yanhui, GAO Meng,et al. Forecasting for non-linear load power based on wavelet and ANN[J].Electrical Measurement & Instrumentation, 2008,45(12): 8-11

[5] YAO Zhisheng, SHAO Chunfu, GAO Yongliang. Research on methods of short-term traffic forecasting based on support vector regression [J].Journal of Beijing Jiaotong University, 2006, 30(3): 19-22.

[6] DENG Naiyang, TIAN Yingyang. A new method in data mining——support vector machine [M].Beijing: Science Publishing Company, 2004.

[7] ZHANG Bing, SU Hongye, CHU Jian. Predictive control algorithm based on least squares support vector machines [J]. Control and Decision, 2004, 19(12): 1399-1403.

[8] HAN Xiaojuan, CAO Hui, LI Yong, et al. Short term wind speed prediction based on wavelet transform and LS-SVM[J]. ACTA ENERGIAE SOLARIS SINICA, 2011, 32(10):1538-1542.

[9] SUN Xiangwen, WANG Meiyu, HUI Jing. Electricity prices predicating based on chaos theory in electric power market [J]. Journal Of Northeast Dianli University Natural Science Edition, 2009, 29 (2): 52-55.

[10] LV Jinhu, LU Junan, CHEN Shihua. Chaotic time series analysis and its application[M].Wuhan, Wuhan University Press, 2002.

[11] XU Lei, ZHANG Fengming, Chen Jun. Research on fault diagnosis method based on PSO neural network [J]. Computer Engineering and Design, 2007, 28(15): 3640-3643

[12] LI Na, LI Yuxia. Electricity price forecast based on PSO-BP neural network [J]. Wuhan University Press, 2008, 41(4):102-105.